Under Review

A potential new approach is being developed for assessing progression-free survival in solid tumour studies using independent review committees

Kristin Borradaile, Kevin Duffy and J Michael O'Neal at BioClinica

Approximately four years ago, an industry committee, the PhRMA PFS Working Group, convened to begin discussing the use of progression-free survival (PFS) endpoint in oncology trials and the value added by independent review facilities. This group was specifically interested in reducing the costs and burden associated with conducting oncology trials with imaging endpoints.

This resulted in the recommendation for a modified independent review approach for large registrational studies. In 2012, the Food and Drug Administration (FDA) referred this discussion to the Oncologic Drugs Advisory Committee (ODAC) for their opinion. In July 2012, ODAC held a public meeting to discuss a potential independent review committee (IRC) approach for the evaluation of PFS in non-haematologic malignancies (1).

ODAC Meeting Summary

The original intent of the IRC was to eliminate the perceived bias associated with investigator assessments. This was thought to be particularly relevant for unblinded studies where the investigator knew which patients were enrolled in the treatment and experimental arms of the trials, or for blinded studies where the toxicity profile associated with treatments had the potential to unblind the treatment assignment.

The implementation and management of the IRC process is costly and burdensome for sites and clinical trial sponsors. The average cost was estimated at $4,500-$7,500 per patient, depending on the size and complexity of the trial. In addition, there is an average discordance rate of 30 per cent between investigator and IRC assessments. This discordance results in informative censoring when the investigator identifies progression and removes the patient from the study without an IRC-identified progression event. Disproportionate informative censoring between treatment arms leads to the introduction of bias into the trial.

Data Analysis

The PhRMA PFS Working Group analysed data from 27 past randomised clinical trials where an IRC was utilised. Despite patient-level discordances, they concluded that there was a high degree of correlation between investigator and IRC-determined PFS treatment effect at the population level. Because FDA drug approval decisions are based on population-level statistics, the use of IRC did not change the outcome of the trials. It was concluded that no systematic bias was introduced by the investigators in the 27 trials. Additional analyses based on trials submitted to the FDA produced similar results.

Based on their analyses, a recommendation was made to utilise IRC in more of a limited role. Specifically, IRC would be used as an audit function in a sample of subjects to evaluate the consistency in treatment effect measured by the investigator sites and IRC. The intent of this approach would be to reduce costs and logistical burden, avoid missing data issues at the IRC, and mitigate statistical issues encountered, such as informative censoring.

Audit Methodology

With the proposed approach, the investigator assessments would be used as the basis for the primary endpoint analysis, instead of the current standard which is to utilise the IRC data. For unblinded trials, an IRC audit of a sample of subjects would be performed to confirm the investigator assessments were free from bias. It was estimated that 100-200 subjects would be required to be read by the IRC as part of the audit; however, the sample size would be variable and dependent on the treatment effect determined by investigator assessments. Based on the required sample size, small studies (phase 1 and phase 2) would require 100 per cent IRC review. For truly blinded trials, the consensus was that IRC review would be desired for the investigators to be aware that a level of oversight was included in the trial.

Two potential audit statistical analyses were proposed:

- Analysis of population-level hazard ratios between the investigator and IRC assessments; or
- Analysis of 'differential discordance' among treatment arms between investigator and IRC assessments

Based on a pre-defined threshold, if bias in investigator assessments was identified, 100 per cent IRC review would

be required. This approach was recommended for solid tumour registrational studies utilising PFS as the primary endpoint. Further discussion would be needed for different tumour types and complex studies that require multiple imaging modalities.

Special Considerations

The data presented during the ODAC meeting was comprised of a 27-trial analysis by the PhRMA PFS Working Group of published data and a similar analysis by the FDA of submitted trial data. Both analyses concluded that no systematic bias had been introduced by the investigators. A high degree of correlation was observed between the investigator- and IRC-determined PFS treatment effects as measured by hazard ratios, and objective response rate as measured by odds ratios.

These analyses may have been conducted using a biased sample of studies. All included trials had independent review performed by an IRC. Therefore, the investigators participating in the trials knew their work would be 100 per cent reviewed by a third party. It is questionable whether their assessments would have differed without this knowledge.

In addition, by including only trials with published data or data submitted to the FDA, it is reasonable to believe that all or most of the trials had favourable results by the IRC. With a larger treatment effect, a high degree of correlation between the investigator and IRC assessments would be expected. It is questionable how the results would have differed if trials with favourable investigator results but unfavourable IRC results, or *vice versa*, had been included.

No Final Decision

While ODAC agreed that the data presented suggests the IRC audit methodology is a reasonable alternative in certain settings, it was noted that additional research and discussion was needed on many points:

- The size of the sample audit may be based on treatment effect observed during the trial
- The method of random subject selection: a truly random sample would likely exclude some investigator sites from the IRC audit
- The timing of random subject selection, if based on treatment effect, could be selected towards the end of the trial
- The statistical method used to compare the investigator and IRC results
- The audit threshold of differential discordance or hazard ratio differences that would result in the need for 100 per cent IRC review
- The timing of the above decision
- The scope of studies eligible for the IRC audit approach. The FDA indicated that the proposed approach applies to solid tumour studies and not to haematologic tumours since other factors, such as blood counts and physical exam findings, are incorporated into the assessment. Regarding solid tumour studies, ODAC indicated further discussion and research was needed in order to make recommendations for tumour types that are more difficult to measure; those which incorporate both radiographic assessments and biomarkers, such as prostate cancer; and those which require additional types of imaging assessments other than CT scans

ODAC agreed that the IRC audit approach is a viable alternative and more cost-effective; however, more research is needed and the specific audit strategy per study needs to be determined on a case-by-case basis with the FDA.

Upcoming Guidance

In August 2011, the FDA released its draft *Guidance for Industry: Standards for Clinical Trial Imaging Endpoints* (2). Within the draft guidance, recommendations are made regarding when centralised image interpretation is important:

"The need for centralised (core) image interpretation process is contingent upon the role of imaging within the trial. In situations where image interpretation results in measurements representing important components of trial eligibility determination or safety or efficacy endpoints, and these measurements are vulnerable to considerable variability among clinical sites, a centralised image interpretation process is needed. A centralised image interpretation process also is critical to controlling bias in open label trials. In general, compared to a site-based image interpretation, the centralised process can better provide verifiable and uniform reader training as well as ongoing management of reader performance, ensuring that the process is accurate and that bias and variability are minimised."

The exception to the above includes, "a randomised, double-blinded clinical trial of an investigational therapeutic drug where the imaging technology is widely available, the image is easily assessed by a

"ODAC agreed that the IRC audit approach is a viable alternative and more cost-effective; however, more research is needed and the specific audit strategy per study needs to be determined on a case-by-case basis with the FDA"

clinical radiologist, and the investigational drug has shown little or no evidence of unblinding effects".

The draft guidance also stresses the importance of outlining detailed information about the blinded review in the Charter, blinded reader qualification, reader training, re-training, and reader performance monitoring. These requirements are logistically difficult to achieve in the site read setting.

It is unlikely that the recent discussion from the ODAC meeting will be incorporated into the final version of this guidance due to its publication date. As a result, it will be necessary for sponsors to seek study-specific guidance and recommendations on the type of independent review approach required.

Shift of Regulatory Burden

Radiology reads performed as part of routine clinical care and those performed by the IRC for the purpose of regulatory approval differ greatly. Radiologists in clinical practice are typically unaware that a patient is enrolled in a clinical trial for which specific protocol requirements apply. Outside of major research institutions, clinical radiologists are largely unfamiliar with the response criteria required in clinical trials and rarely select and measure target lesions at each time point. In addition, it is not uncommon for a different radiologist to read each time point for an enrolled patient, increasing the variability of assessments. In order to meet clinical trial requirements, a high percentage of trial sites utilise study coordinators or oncologists to measure the lesions that are qualitatively referenced in radiology reports in order to fulfil the case report form (CRF) requirements associated with clinical trial protocols.

The draft *Guidance for Industry* goes into great detail regarding the burden of reducing variability in assessments during imaging trials. This includes specifications regarding image quality assessments, qualification and training of readers, standardisation of image display and interpretation, the measurement tools and reading system, options and requirements for image manipulation, and monitoring of reader variability and assessments.

The need to make site reads auditable, as they will comprise the primary endpoint data, simply shifts the cost and burden from the IRC to the sites and is logistically infeasible. The following considerations apply:

- Cost and training of radiologists
- Training of 100+ sites when some may only enrol a couple of subjects
- Ensuring site compliance as it relates to image quality and adherence to protocol requirements
- Difficulty and cost of mandating standardisation
- Management and monitoring of the reader process
- Process for replacing readers whose performance falls outside of acceptable thresholds

The IRC workflow is specifically designed to produce greater consistency and reduced variability in image interpretation. The main components of the IRC workflow are:

- Consistent small group of radiologists
- Standardised reader training, qualification and testing
- Controlled image viewing
- Standardised measurement tools
- Derivation procedures and edit checks to ensure response criteria are applied appropriately and consistently
- Variability monitoring
- Fully auditable source data

Audit Concerns

There are concerns regarding the timing of IRC audit analysis decisions as it relates to the timing of regulatory filings. Depending on when the decision is made regarding whether a 100 per cent IRC review is needed, there could be a delay in the sponsor's filing, resulting in increased costs for the sponsor and a delay in access to treatment for patients that need it. The potential that the FDA could retrospectively request a 100 per cent IRC review during the time of new drug application review further increases this concern.

Further concerns relate to the auditability of site reads, especially if they will be used as the primary endpoint in regulatory submissions. The source data in imaging trials is the image with the lesion measurements and annotations included. The data transcribed onto trial CRFs do not represent source documentation. CRFs do not typically contain screen shots of annotated images and annotations are not typically saved on clinical review systems at the sites. Therefore, monitoring typically occurs between radiology reports and CRFs, not the annotated images. Conversely, the IRC is able to provide the FDA with annotated image archives demonstrating all tumour assessments contributing to trial outcomes.

In addition, with the proposed audit approach, statistical issues can arise when 100 per cent IRC review is retrospectively required. Currently, sponsors employ a reading paradigm where the IRC conducts real-time progression confirmation reads to prevent statistical issues associated with informative censoring. If an IRC audit approach is implemented, and 100 per cent IRC review is later determined to be necessary, informative censoring could become an issue since real-time confirmation reads would not have been prospectively performed.

Recommended Approach

With the need for an IRC audit of a random sample of subjects and the potential need to conduct 100 per cent IRC review, it will be necessary for the IRC to collect all subject images throughout the trial in a collect-and-hold scenario. Random subject selection will occur towards the latter part of the study. To mitigate unnecessary delays and the potential for missing image data, it is recommended that the image collection occurs upfront.

A primary concern will be the potential for the FDA to require 100 per cent IRC review following the completion of the IRC audit. Depending on the timing of this decision, significant delays could affect the timing of the regulatory filing and review.

Another risk mitigation strategy is an upfront IRC audit based on the first subjects enrolled into the trial. While these may not be the same subjects randomly selected for the FDA-required IRC audit, an early analysis of the investigator/IRC assessments of the first enrolled subjects could provide more information as to whether or not it is likely that a full IRC review would be required. If the initial results suggest investigator bias is present, the sponsor may opt to have the IRC read 100 per cent of scans prospectively to mitigate the potential delay associated with a late decision to require 100 per cent review.

References

1. **FDA Briefing Document – Oncologic Drugs Advisory Committee Meeting: Evaluation of radiologic review of progression-free survival in non-hematologic malignancies, 24th July 2012**
2. ***Draft Guidance for Industry: Standards for clinical trial imaging endpoints*, FDA, August 2011**

About the authors

Kristin Borradaile is the Director of Clinical Operations at BioClinica where she is responsible for developing operational strategies and overseeing the global project management teams. Kristin holds an MBA from Rider University, an MS in Quality Assurance and Regulatory Affairs from Temple University, and a BSc in Biology from the University of Pittsburgh.

Email: kristin.borradaile@bioclinica.com

As BioClinica's Senior Vice President, Global Clinical Operations, **Kevin Duffy** is the Corporate Auditor of the Japan K.K. and directs strategic initiatives and day-to-day business activity. Kevin has over 27 years of experience in the healthcare industry. He holds an MBA from Pepperdine University and a BA in Economics from the University of Michigan.

Email: kevin.duffy@bioclinica.com

J Michael O'Neal, BioClinica's Chief Medical Officer, provides leadership and oversight to staff radiologists and external radiology consultants. Dr O'Neal joined BioClinica in 2004 as a diagnostic radiologist with more than 20 years of practice experience. He is certified by the American Board of Radiology and has been published in *The Journal of Bone and Joint Surgery*.

Email: michael.oneal@bioclinica.com